Identification of Major Histocompatibility Complex Haplotypes in Icelandic Horses

Honors Thesis
Presented to the College of Agriculture and Life Sciences, Animal Science
of Cornell University
in Partial Fulfillment of the Requirements for the
Research Honors Program

by
Nathaniel Violette

May 2014

Dr. Douglas F. Antczak

Abstract

The Major Histocompatibility Complex (MHC) is a region of the vertebrate genome that encodes many immune-related proteins. The most well-known proteins that are encoded in this region are the MHC class I and MHC class II cell surface proteins. These proteins are integral in producing the adaptive immune response. Using polymorphic microsatellites within the MHC, MHC haplotypes were characterized from an extended Icelandic horse family of half and full siblings. The herd is composed of one stallion, fifteen dams, and three sets of foals from each dam. A total of nine microsatellite loci were tested: two in the MHC class I region, two in the MHC class III region, and five in the MHC class II region. The Icelandic herd displayed a large degree of diversity in this region even though the population has been isolated and affected by bottleneck events in the past. From the 32 possible unique haplotypes belonging to the one founder stallion and 15 mares, 26 unique haplotypes were identified. A MHC recombination event occurred on one of the paternal chromosomes. The stallion also exhibited segregation distortion in the distribution of the stallion's MHC haplotypes inherited by the foals but not the sex ratio of the offspring. These findings suggest that the Icelandic horse population has a greater degree of diversity in the MHC region than most other commonly studied horse breeds. This diversity may come from a recombination mechanism that produces unique haplotypes by crossing over of existing haplotypes in the population at a higher rate than found in larger populations. Another possible explanation would be that less favorable MHC haplotypes are allowed to persist in the less immunologically selective environment of Iceland.

Acknowledgements

First and foremost, I would like to thank Dr. Douglas F. Antczak for giving me the opportunity to pursue my research interests in genetics as well as equine health. I am grateful for his support of my work and encouragement of my ideas throughout this project. Dr. Antczak always provided clear answers to all of my questions. He also was always able to ask the right questions of me that stimulated my critical thinking and kept developing my interest in research.

I would also like to thank Don Miller for his enormous amount of help and support throughout my lab training and all of my lab work. I want to thank Don for being extremely reliable and flexible for the past year and a half that I have spent in the lab. His relaxed attitude as well as his extensive knowledge of genetics and molecular techniques made for a wonderful experience working in the laboratory. I want to thank Becky Harman and Peggy Brosnahan for their expert assistance whenever Don was not in the lab.

I wish to thank Dr. Bettina Wagner for her guidance with my questions about equine immune system function. Dr. Wagner was very generous to let me work with her herd of Icelandic horses. I also want to thank Alison Keggan and other members of the Wagner lab for their patience collecting blood samples from the Icelandic herd and for sending me pictures of the foals.

I want to thank my advisor, Dr. James Giles, for his support throughout my years at Cornell. He has provided me with great advice and his optimism has always helped me through stressful times.

The Antczak laboratory along with the James A. Baker Institute for Animal Health has been a perfect setting to begin my research experience. I thank all of the people at the institute for creating a supportive and friendly atmosphere that has been a key to success for this project.

I want to thank all of my friends and family for encouraging me to pursue my interest in research and supporting me throughout the process.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Introduction

The Major Histocompatibility Complex (MHC) is a region of the vertebrate genome that contains the genes that are transcribed and translated to produce most of the organism's immune proteins. These molecules mediate the interactions between immune cells, body cells, and pathogens. The most well-known and prominent of these proteins are the MHC class I and class II cell surface proteins (Abbas, 2000). MHC class I proteins are presented on all healthy cells of the body and act to help the body differentiate healthy cells from infected or cancerous cells. MHC class II surface proteins are used by immune cells, more specifically T lymphocytes, to recognize foreign particles such as proteins or antigens. The MHC is organized into three class regions; they are found in the order class I, class III, and class II on the chromosome.

Unsurprisingly, most research of the MHC has been done on humans. The human MHC has the highest gene density as well as the largest degree of polymorphism in the genome. The recombination rate for the region is lower than the average recombination rate of the entire genome. Researchers have found that recombination events in the MHC occurred at localized regions while other regions are highly conserved (Lam *et al.* 2013*).* Recombination is an important biological process which drives genetic polymorphism and genome evolution. Certainly, the mechanism behind recombination of the MHC serves an essential role in effective immune response not only in humans but all vertebrates. The MHC can be defined using molecular genetic techniques to make haplotypes. Although, identification of MHC haplotypes is a difficult task due to the high levels of variation and linkage disequilibrium within this

region (Traherne, 2008). **Figure 1.1** shows the region of the human chromosome that contains the MHC, the structure of the MHC and how the genes located in the region regulate important immune function. This region is located on chromosome 6 in humans. Both human and equine chromosomes containing the MHC are very similar and can be shown interchangeably in the figure below.

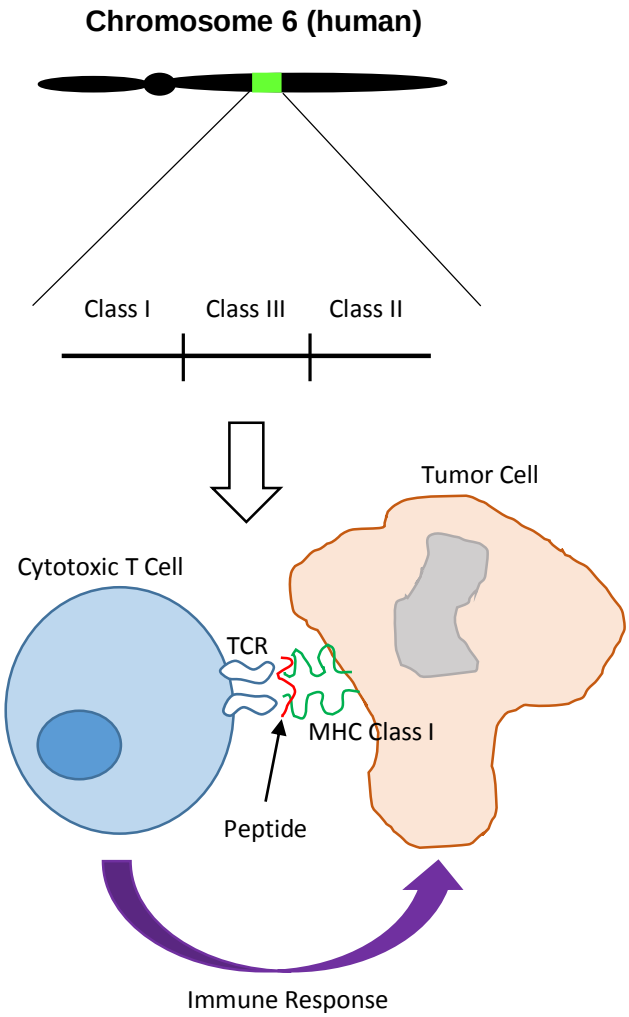

Figure 1.1 MHC Function Diagram shows the chromosome containing the MHC in humans, the organization of the class regions, and an example of how these genes produce proteins that are used in the adaptive immune system. Figure adopted from (Abbas, 2000).

Microsatellites are regions of the genome in which two, three, or even four bases repeat many times. These regions can be very polymorphic due to errors made by DNA polymerase in replication. The polymerase infrequently stutters in these regions and adds on extra repeats. These extra repeats are inherited by that animal's offspring (Tozaki, 2001). Therefore, these regions can be amplified by PCR and their lengths determined. Compiling the allele lengths in a certain region can create haplotypes of that region if relatives of that animal have been tested. **Figure 1.2** shows two examples of microsatellite alleles.

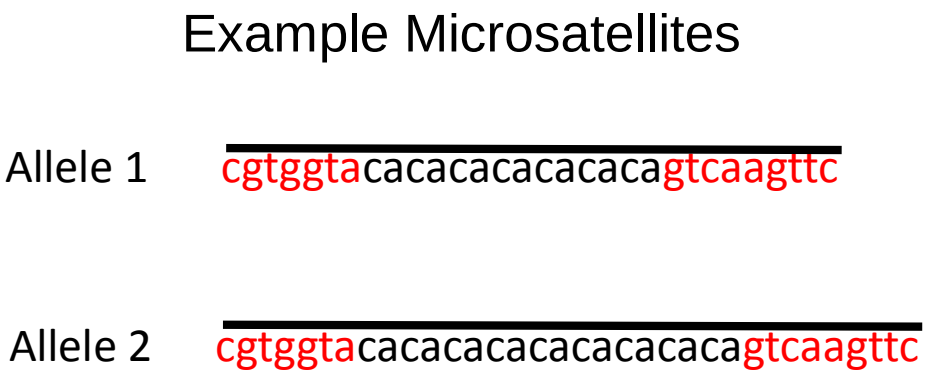

Figure 1.2 Microsatellite Diagram. This example shows a (ca) dinucleotide repeat in which allele two has had two extra repeats added on. Therefore, allele two is four bases longer. The red regions are variable flanking regions. Primers can be made to anneal to these regions in order to amplify the region.

1.2 Literature Review

The MHC in the horse is located on chromosome 20 and is also named the Equine Leukocyte Antigen (ELA). A bacterial artificial chromosome (BAC) contig has been made up of all of the genes in the equine MHC (Fraser *et al.* 1996*)*. The class II region codes for the proteins that make up the class II surface proteins. There are two distinct regions within class II, DR and DQ. The DQ region codes for the actual proteins and the DR region is the upstream regulatory region (URR) for DQ. The ELA-DRB1 region codes for the regulatory factor binding site that regulates expression of DQ genes. The DQ gene is highly conserved (Diaz *et al.* 2005*)*. In contrast, many regions of the MHC are highly polymorphic which generates the genetic diversity needed for the immune system to target a huge range of pathogen. Studies have shown that less genetic diversity in the equine MHC correlates with susceptibility to autoimmune disease (Klumplerova, *et al.* 2013*)*

The purpose of this investigation was to continue research on the MHC region of the equine genome as well as look into the diversity seen in the Icelandic horse breed. Many papers have been published on the use of microsatellites on the equine MHC. Many polymorphic microsatellite locations have been determined, for example a certain study found 37 of these locations in the equine MHC (Brinkmeyer-Langford *et al.* 2013). This paper used these microsatellites as well as published SNPs to run linkage disequilibrium (LD) analysis of the region. They found a large amount of LD which indicates structural polymorphisms and past recombination events. Microsatellite markers are reliable and efficient tools in characterizing haplotypes in the MHC region. The MHC genes are co-dominantly expressed, this means that each individual is able to

express all alleles on both parental chromosomes. The lengths of these fragments are then determined and used to create haplotypes for each of the animals.

Tseng *et al.* 2010 used five polymorphic microsatellites in the equine MHC to characterize haplotypes in 353 horses of a variety of breeds and found 50 haplotypes. These haplotypes were compared with established haplotypes and had 12 in common. Also, a number of microsatellite subhaplotypes were associated with five equine leucocyte antigen serotypes which demonstrates a linkage to the MHC class I protein polymorphisms. The results of this paper validate the effectiveness of microsatellite markers at characterizing MHC haplotypes.

The first genetic studies that involve the Icelandic horse breed were done during the late 1970's. Adalsteinsson, 1978 describes the inheritance of coat color in the Icelandic breed. Icelandic horses are known for their variety of coat colors. These studies concluded that the horses had a high degree of heterozygosity in their coat color genes which allows for the wide range of observed phenotypes.

The population of horses in Iceland are a unique subgroup due to their isolation and small founding population. The Icelandic horses have been under control of natural and artificial selection since settlers brought them to Iceland around 900AD. The population has been purebred for more than 1000 years due to laws passed by the Icelandic government (Campana *et al.* 2001). The breed has also experienced extreme bottlenecking events such as the 1783 eruption of Lakagigar which lead to the death of about 70% of the Icelandic horses (Hendricks, 1995). These horses are also isolated from many of the pathogens found commonly in other horse populations around the world. For this reason, Icelandic horses are brought to other countries and their

reaction to natural pathogens found in different areas is studied. These circumstances make the genetic diversity of this breed's MHC a topic of interest.

A population of Indian rhesus macaques have been studied for MHC diversity and recombination events. The researchers found that a subpopulation of these macaques, that became isolated and bottlenecked, have shuffling of conserved segments of their MHC generating more haplotypes (Doxiadis *et al.* 2013). This mechanism of developing MHC diversity through recombination in a small founder population may be an interesting parallel between Indian rhesus macaques and the Icelandic horse breed.

1.3 Aim of this Investigation

The goal of this project was to identify haplotypes of the equine MHC using nine microsatellites found throughout the MHC in a herd of Icelandic horses at the Cornell College of Veterinary Medicine. This thesis was meant to estimate the degree of genetic variation and polymorphism found in the MHC of the Icelandic horse population.

Chapter 2: Materials and Methods

2.1 Horses

The 15 mares and 1 stallion used in this study were imported from Iceland about 3 years ago and are now kept by Cornell University Veterinary School. The mares and stallion produced the first crop of fifteen horses while still in Iceland. Since then, two breeding seasons have passed and two crops of foals have been produced at Cornell University creating a herd of 61 total horses (**Table 2.1**).

Table 2.1 Summary of Icelandic Horse herd

Horse	Accession #	Number of Horses
Stallion	4279	1
Mares	4264-4278	15
Crop 1 (2011)	4298-4312	15 (6 female, 9 male)
Crop 2 (2012)	4313-4327	15 (5 female, 10 male)
Crop 3 (2013)	4347-4354	15 (9 female, 6 male)

2.2 Microsatellites

The PCR primers that anneal to certain microsatellite regions in the MHC are organized in **Table 2.2**. These primers are fluorescently labeled so that amplified fragments can be detected during fragment analysis. The locations of the microsatellites are shown in **Figure 2.1**.

Table 2.2 List of MHC PCR Primers. The labels NED and FAM are fluorescent dyes produced by Applied Biosystems.

Name	MHC Class Region	Sequence	Product Size (bp)	Label
Cor110_F	1	TTTGGTCTTTGCAGGTATGG	193-225	FAM
Cor110_R		TCTCCCTTCCTCTTTGTTCC		
TAMU_305_93_F	1	GAAGCCCAGTCTGAGTGAAGAT	343-347	FAM
TAMU_305_93_R		AGATTTGGACCGAGAAAGTCTG		
ABGe_9019_F	3	CTGAGAGAGACAGCATTTGTGG	299-313	NED
ABGe_9019_R		GAAAGGTGTCTCCATTGTTGCT		
ABGe_9030_F	2	CCAGCAGACCTGCAAGAGTA	205-219	NED
ABGe_9030_R		AGCATGAGAGCCATGAAGGT		
TKY3324_F	2	AGCCGTCCTGTTCCAACTAA	255-273	FAM
TKY3324_R		TGCCCCTTAAAACTCTGTCTTT		
Cor112_F	2	TTACCTGGTTATTGGTTATTTGG	236-268	NED
Cor112_R		TCACCCACTAAATCTCAAATCC		
Cor113_F	2	TGTTTAGAACTCGCCAGGAG	260-280	NED
Cor113_R		TCATCAGTTCCTTGCCTAGC		
UM011_F	2	TGAAAGTAGAAAGGGATGTGG	160-184	NED
UM011_R		TCTCAGAGCAGAAGTCCCTG		
Cor114_F	2	TCAAAATCCACACTCCCTTC	235-255	FAM
Cor114_R		TCCATAAAGAGTGGGACACTG		

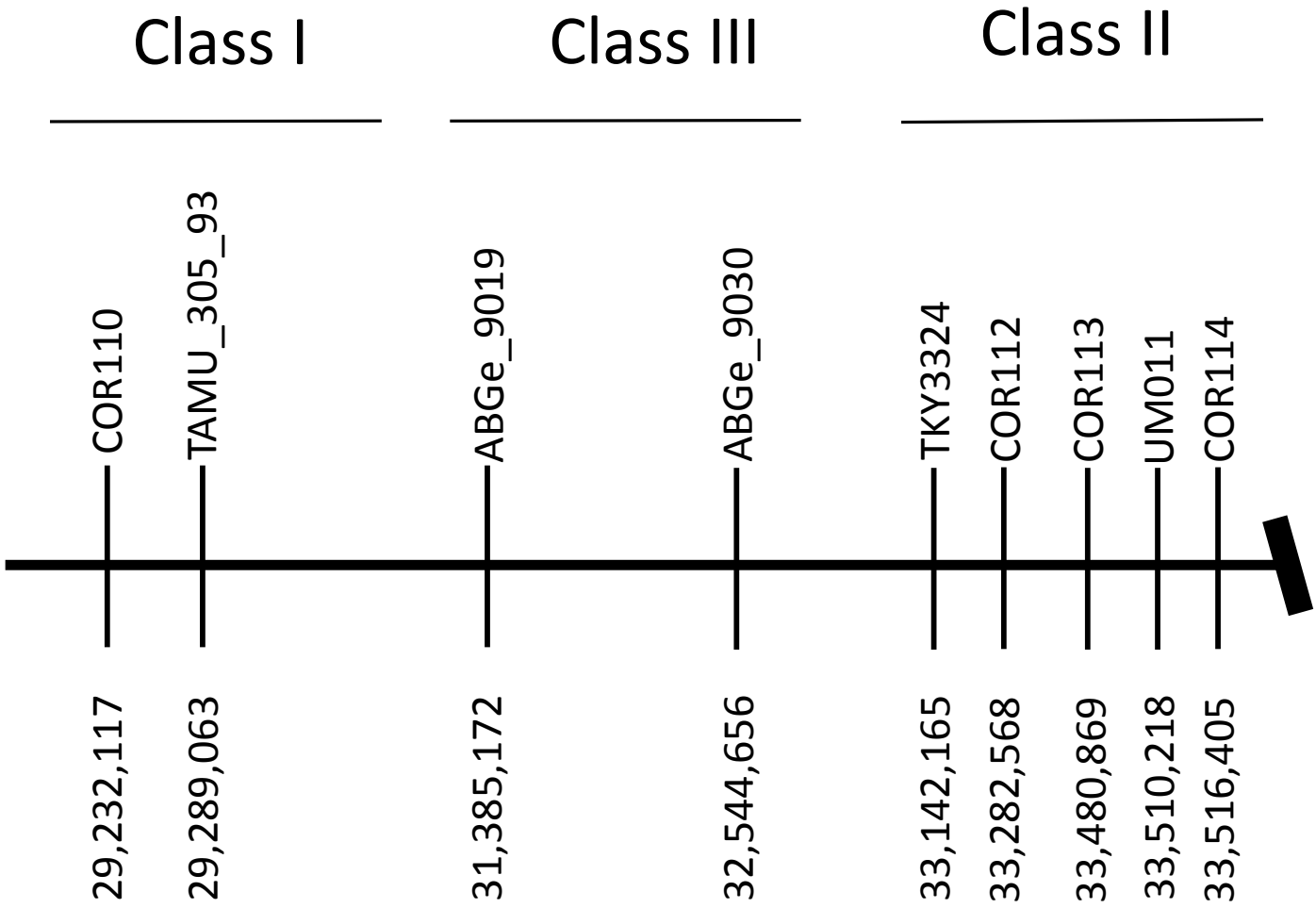

Figure 2.1 Map of Major Histocompatibility Complex microsatellite loci. Microsatellite base pair location on equine chromosome 20 indicated.

2.3 Horse Blood Samples

Blood samples from the 15 foals in Iceland were sent to Cornell and blood samples were taken from the 15 mares and stallion upon arriving at Cornell. Samples were obtained from each new crop of foals every year. All samples were taken following the guidelines provided by Cornell's Institutional Animal Care and Use Committee (IACUC) protocol #1986-0216. Blood samples were collected using heparinized tubes from the jugular vein.

2.4 Isolation of Genomic DNA from Peripheral Blood

Blood samples were stored at 20°C until the peripheral blood lymphocytes were isolated. Genomic DNA (gDNA) was isolated from the samples of blood using Qiagen kit (DNeasy Blood & Tissue Kit (250) Cat. No. 69506) following the manufacturer's protocol (Qiagen Inc., Valencia CA).

2.5 Selection of Microsatellite Markers

The intra-MHC markers; Cor110, Cor112, Cor113, UM011, and Cor114 were selected due to their performance in previous experiments (Tseng *et al.* 2010). These five were previously found to be polymorphic, stable and accurate. Four additional intra-MHC markers were taken from a panel created by Texas A&M University (Brinkmeyer-Langford *et al.* 2013); TAMU_305_93, ABGe_9019, ABGe_9030, TKY3324. These primers were chosen based on their polymorphism, class location in the equine MHC and allele lengths. Assays for the additional four were optimized at Cornell University and added to the study.

2.6 Equine MHC Microsatellite Amplification

The polymerase chain reaction (PCR) used to amplify microsatellite alleles consisted of 50 ng of gDNA (2 µl) added to a 10 µl reaction mix. This mix contained 10x PCR Buffer (1 µl), 2.5 µM dATP, dCTP, dTTP, dGTP (0.8 µl combined), 50 µM $MgCl_2$ (0.3 µl) 1 U *Taq* polymerase (0.2 µl),10 µM fluorescent forward and reverse primers (0.25 µl each) (**Table 2.2**), and 5.2 µl of water. In some cases, multiplexes were used to amplify two markers in one reaction. Multiplexes of intra-MHC markers, TAMU_305_93 and ABGe9019 and ABGe9030 and TKY3324 were used as pairs, respectively. The first pair used identical concentrations and volumes of reagents in the previously described reactions and the extra volume taken up from the second pair of primers was taken from the water added to the reaction. The primer volumes for the second pair were offset with 0.125 µl and 0.375 µl for each primer, respectively. This adjustment optimized the reaction. The same amplification protocol was used for each reaction; 1 cycle of 95°C for 5 minutes, followed by 33 cycles of 95°C for 1 minute, 58°C for 1 minute, and 72°C for 30 seconds, and 1 cycle of 72°C for 20 minutes for the final extension.

2.7 Fragment Analysis

To analyze the microsatellite fragments, 2 µl of PCR product was mixed with 1 µl of GeneScan™-500 LIZ® Size Standard and 14 µl of Hi-Di™ Formamide (Applied Biosystems) to produce a final volume of 17 µl. The samples were

electrophoresed on an ABI3700 (Applied Biosystems DNA Analyzer) at the Cornell University DNA core facility, Ithaca, NY. Fragment lengths were analyzed using Applied Biosystems PeakScanner v1.0 software. **Figure 2.2** shows an electropherogram from the PeakScanner program which displays a heterozygous horse having fragment lengths of 168bp and 172bp. By testing trios of stallion, mare, and foal, it was possible to phase the microsatellite markers and determine MHC haplotypes for all individuals.

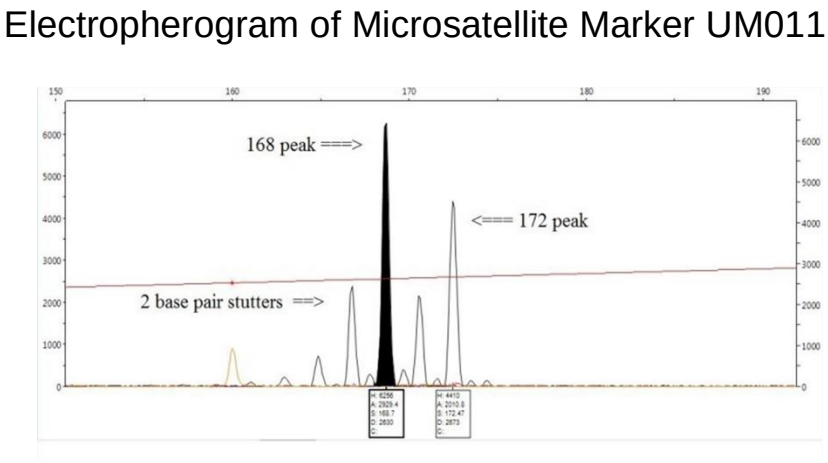

Figure 2.2 Screenshot of the PeakScanner software. Shows two peaks indicating fragment lengths of 168 and 172bp.

2.8 Calculation of Segregation Distortion

A chi-squared test was applied to MHC haplotype segregation and sex ratio segregation in the sire, assuming even segregation using the statistical software R (Bell Laboratories, http://www.r-project.org/).

Chapter 3: Results

The data collected in this study was produced exclusively using microsatellite marker fragment lengths. The region is mapped out in **Figure 2.1**. By identifying microsatellite lengths from a family, including the stallion, mare and foal, microsatellite haplotypes were easily characterized based on inheritance of MHC haplotypes by the foal from the parents. Since three foals were produced by each mare, the microsatellite lengths can be further confirmed. Phasing the microsatellite lengths allows for logical deduction of MHC haplotypes, an example of this process is illustrated in **Table 3.1** and **Table 3.2**. These MHC microsatellite haplotypes translate into the MHC alleles of each horse. Assigning specific MHC alleles to each horse allows the degree of MHC diversity in this herd to be determined.

3.1 Haplotype Identification

The herd of the original fifteen mares and stallion contained 32 possible unique haplotypes (32 chromosomes) and of those possibilities, 26 unique haplotypes were found. The haplotype 24 was a new haplotype formed by recombination found in one of the foals. The microsatellite haplotypes along with their unique color and allele lengths can be found in **Table 3.3**. The colors from this figure correspond with the colors found on the lineage web (**Figure 3.1**). This web shows the stallion as the large rectangle in the center with his accession number and his MHC haplotypes signified as the yellow (Haplotype 1) and light blue (Haplotype 2) sides of the rectangle. The ring of ovals around the stallion signify the mares with their accession numbers and MHC haplotypes as colors. The outermost layer are the foals which are both rectangles (male) and ovals

(female) with their inherited haplotypes as colors. The foals that were born the first season are innermost and radiating out are the next two consecutive breeding seasons. Each of these groups of horses are labeled by family numbers; 1-15. The haplotype colors and numbers in each family can be viewed in **Table 3.4**. All of the fragment lengths for each horse and their families can be found in **Table 3.5**.

Example of Haplotype Phasing

Table 3.1 Unphased Fragment Lengths. Arrows indicate which lengths were switched to align with the correct haplotype. Each of these switches are made based on the theory that each fragment length is attached to one chromosome and must be transferred with all the other microsatellites on that chromosome (except in the case of recombination).

	Hap	Cor110	Cor112	Cor113	UM011	Cor114
Sire, 4279		197	248	270	172	245
		225	254	270	184	249
Dam, 4264		207	262	270	172	237
		217	264	280	172	247
Foal 1, 4298		197	254	270	172	237
		217	262	270	172	249
Foal 2, 4313		217	248	270	172	237
		225	262	270	184	245

Table 3.2 Phased Haplotypes. The family below showing how haplotypes are aligned after phasing and how they are transferred from parents to offspring.

	Hap	Cor110	Cor112	Cor113	UM011	Cor114
Sire, 4279	1	197	254	270	172	249
	2	225	248	270	184	245
Dam, 4264	3	207	264	280	172	247
	4	217	262	270	172	237
Foal 1, 4298	1	197	254	270	172	249
	4	217	262	270	172	237
Foal 2, 4313	4	217	262	270	172	237
	2	225	248	270	184	245

Table 3.3 MHC haplotypes found in the Icelandic horse herd. Each of the fragment lengths that make up a haplotype are shown under each microsatellite marker. The number of occurrences found with each of the haplotypes in the 15 mares and 1 stallion is shown (# of chrs). And the number of alleles at each locus is shown in the bottom row.

Haplotype	# of chrs	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
1	1	197	345	301	219	255	254	270	172	249
2	2	225	345	303	205	259	248	270	184	245
3	2	207	345	301	215	255	264	280	172	247
4	2	217	341	299	215	273	262	270	172	237
5	2	209	345	301	205	259	248	270	184	245
6	1	215	341	303	205	259	254	270	172	249
7	1	205	345	309	211	263	256	268	166	247
8	1	205	345	313	207	263	256	278	168	247
9	2	211	345	303	213	259	244	272	168	255
10	1	211	345	303	205	269	262	268	178	251
11	1	207	345	307	211	269	260	260	176	241
12	1	217	341	299	211	257	260	260	176	241
13	1	221	345	301	211	259	248	270	184	245
14	1	209	347	311	209	263	260	270	172	249
15	1	209	345	301	205	259	248	270	172	249
16	1	209	345	311	211	269	260	260	178	241
17	1	211	345	303	213	259	244	270	172	249
18	1	211	345	313	211	259	264	270	184	245
19	1	209	345	301	207	263	256	260	172	243
20	1	205	345	313	207	255	268	264	170	239
21	1	211	343	303	205	263	256	260	172	243
22	1	205	345	311	211	263	260	260	176	241
23	1	223	345	301	207	267	264	270	172	249
24	1	197	345	301	205	259	248	270	184	245
25	1	207	345	301	207	255	264	280	172	247
26	2	209	345	303	205	259	248	270	184	245
27	1	207	345	307	211	269	260	260	178	241
Number of Alleles =>		10	4	7	7	7	8	7	7	9

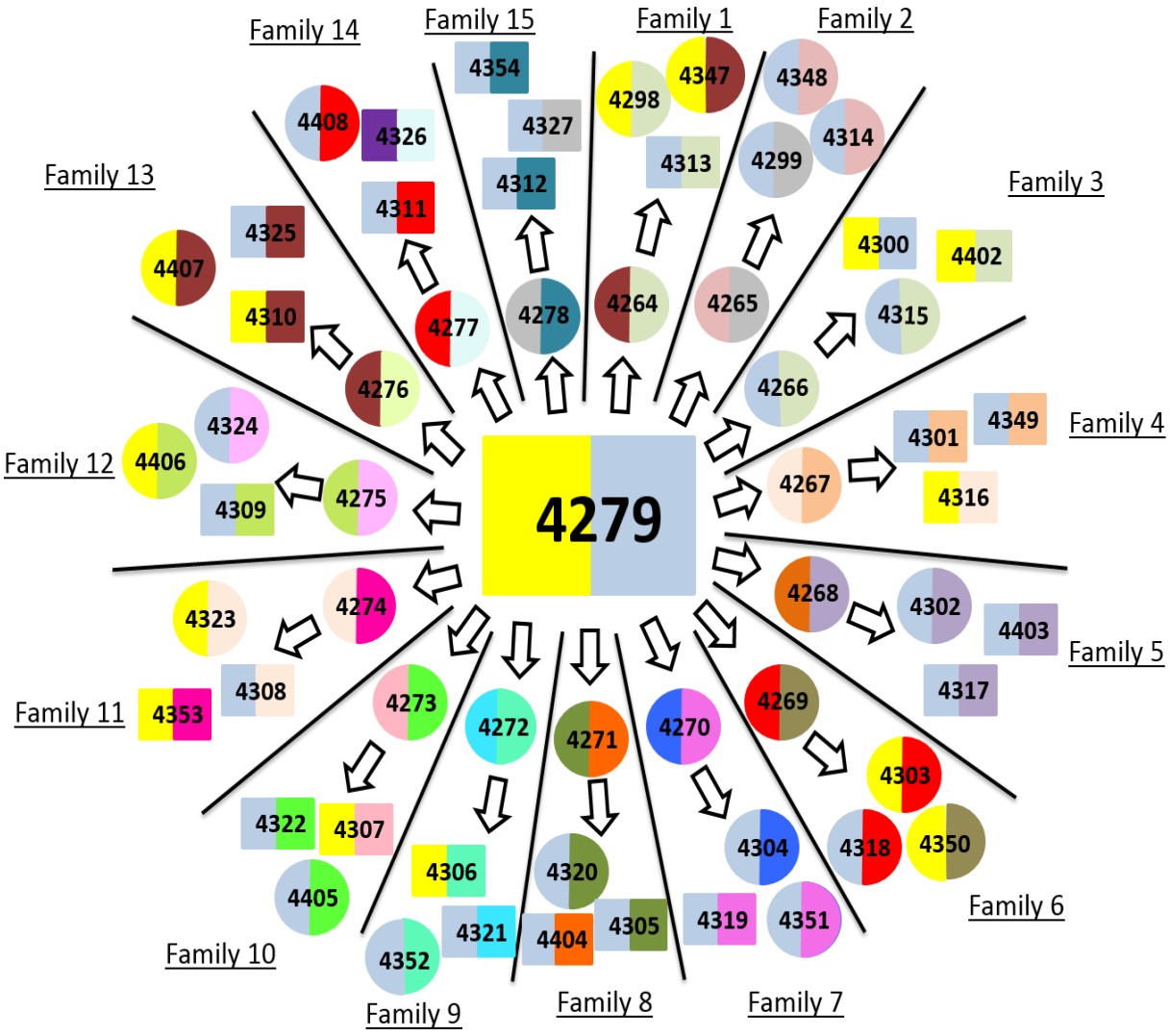

Figure 3.1 Lineage web of herd, showing all horses with ID numbers and haplotypes. The horse in the center (4279) is the sire surrounded by the mares, which are surrounded by the foals. All of the horses' haplotypes are signified by distinct colors.

Table 3.4 Families of horses and haplotypes shown by color and number. Hap1 in the foals is the paternal haplotype and Hap2 is the maternal haplotype. Foals in pink are female and foals in blue are male.

Horse	Acc.#	Hap1	Hap2
Stallion	4279	1	2

Family	Horse	Acc.#	Hap1	Hap2
1	Mare	4264	3	4
	F1	4298	1	4
	F2	4313	2	4
	F3	4347	1	3
2	Mare	4265	5	6
	F1	4299	2	6
	F2	4314	2	5
	F3	4348	2	5
3	Mare	4266	2	4
	F1	4300	1	2
	F2	4315	2	4
	F3	4402	1	4
4	Mare	4267	7	26
	F1	4301	2	7
	F2	4316	1	26
	F3	4349	2	7
5	Mare	4268	8	25
	F1	4302	2	8
	F2	4317	2	8
	F3	4403	2	8
6	Mare	4269	9	10
	F1	4303	1	9
	F2	4318	2	9
	F3	4350	1	10
7	Mare	4270	11	12
	F1	4304	2	11
	F2	4319	2	12
	F3	4351	2	12

Family	Horse	Acc.#	Hap1	Hap2
8	Mare	4271	13	14
	F1	4305	2	13
	F2	4320	2	13
	F3	4404	2	14
9	Mare	4272	15	16
	F1	4306	1	16
	F2	4321	2	15
	F3	4352	2	15
10	Mare	4273	17	18
	F1	4307	1	17
	F2	4322	2	18
	F3	4405	2	18
11	Mare	4274	5	19
	F1	4308	2	5
	F2	4323	1	5
	F3	4353	1	19
12	Mare	4275	20	21
	F1	4309	2	20
	F2	4324	2	21
	F3	4406	1	20
13	Mare	4276	3	22
	F1	4310	1	3
	F2	4325	2	3
	F3	4407	1	3
14	Mare	4277	9	23
	F1	4311	2	9
	F2	4326*	24	23
	F3	4408	2	9
15	Mare	4278	5	27
	F1	4312	2	27
	F2	4327	2	5
	F3	4354	2	27

*Foal with MHC recombination

Table 3.5 Families of horses and all of their phased microsatellite lengths

Family 1	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4264	3	207	345	301	215	255	264	280	172	247
	4	217	341	299	215	273	262	270	172	237
f1_4298	1	197	345	301	219	255	254	270	172	249
	4	217	341	299	215	273	262	270	172	237
f2_4313	2	225	345	303	205	259	248	270	184	245
	4	217	341	299	215	273	262	270	172	237
f3_4347	1	197	345	301	219	255	254	270	172	249
	3	207	345	301	215	255	264	280	172	247

Family 2	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4265	5	209	345	301	205	259	248	270	184	245
	6	215	341	303	205	259	254	270	172	249
f1_4299	2	225	345	303	205	259	248	270	184	245
	6	215	341	303	205	259	254	270	172	249
f2_4314	2	225	345	303	205	259	248	270	184	245
	5	209	345	301	205	259	248	270	184	245
f3_4348	2	225	345	303	205	259	248	270	184	245
	5	209	345	301	205	259	248	270	184	245

Family 3	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4266	2	225	345	303	205	259	248	270	184	245
	4	217	341	299	215	273	262	270	172	237
f1_4300	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
f2_4315	2	225	345	303	205	259	248	270	184	245
	4	217	341	299	215	273	262	270	172	237
f3_4402	1	197	345	301	219	255	254	270	172	249
	4	217	341	299	215	273	262	270	172	237

Family 4	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4267	26	209	345	303	205	259	248	270	184	245
	7	205	345	309	211	263	256	268	166	247
f1_4301	2	225	345	303	205	259	248	270	184	245
	7	205	345	309	211	263	256	268	166	247
f2_4316	1	197	345	301	219	255	254	270	172	249
	26	209	345	303	205	259	248	270	184	245
f3_4349	2	225	345	303	205	259	248	270	184	245
	7	205	345	309	211	263	256	268	166	247

Family 5	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4268	25	207	345	301	207	255	264	280	172	247
	8	205	345	313	207	263	256	278	168	247
f1_4302	2	225	345	303	205	259	248	270	184	245
	8	205	345	313	207	263	256	278	168	247
f2_4317	2	225	345	303	205	259	248	270	184	245
	8	205	345	313	207	263	256	278	168	247
f3_4403	2	225	345	303	205	259	248	270	184	245
	8	205	345	313	207	263	256	278	168	247

Family 6	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4269	9	211	345	303	213	259	244	272	168	255
	10	211	345	303	205	269	262	268	178	251
f1_4303	1	197	345	301	219	255	254	270	172	249
	9	211	345	303	213	259	244	272	168	255
f2_4318	2	225	345	303	205	259	248	270	184	245
	9	211	345	303	213	259	244	272	168	255
f3_4350	1	197	345	301	219	255	254	270	172	249
	10	211	345	303	205	269	262	268	178	251

Family 7	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4270	11	207	345	307	211	269	260	260	176	241
	12	217	341	299	211	257	260	260	176	241
f1_4304	2	225	345	303	205	259	248	270	184	245
	11	207	345	307	211	269	260	260	176	241
f2_4319	2	225	345	303	205	259	248	270	184	245
	12	217	341	299	211	257	260	260	176	241
f3_4351	2	225	345	303	205	259	248	270	184	245
	12	217	341	299	211	257	260	260	176	241

Family 8	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4271	13	221	345	301	211	259	248	270	184	245
	14	209	347	311	209	263	260	270	172	249
f1_4305	2	225	345	303	205	259	248	270	184	245
	13	221	345	301	211	259	248	270	184	245
f2_4320	2	225	345	303	205	259	248	270	184	245
	13	221	345	301	211	259	248	270	184	245
f3_4404	2	225	345	303	205	259	248	270	184	245
	14	209	347	311	209	263	260	270	172	249

Family 9	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4272	15	209	345	301	205	259	248	270	172	249
	16	209	345	311	211	269	260	260	178	241
f1_4306	1	197	345	301	219	255	254	270	172	249
	16	209	345	311	211	269	260	260	178	241
f2_4321	2	225	345	303	205	259	248	270	184	245
	15	209	345	301	205	259	248	270	172	249
f3_4352	2	225	345	303	205	259	248	270	184	245
	15	209	345	301	205	259	248	270	172	249

Family 10	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4273	17	211	345	303	213	259	244	270	172	249
	18	211	345	313	211	259	264	270	184	245
f1_4307	1	197	345	301	219	255	254	270	172	249
	17	211	345	303	213	259	244	270	172	249
f2_4322	2	225	345	303	205	259	248	270	184	245
	18	211	345	313	211	259	264	270	184	245
f3_4405	2	225	345	303	205	259	248	270	184	245
	18	211	345	313	211	259	264	270	184	245

Family 11	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4274	26	209	345	303	205	259	248	270	184	245
	19	209	345	301	207	263	256	260	172	243
f1_4308	2	225	345	303	205	259	248	270	184	245
	26	209	345	303	205	259	248	270	184	245
f2_4323	1	197	345	301	219	255	254	270	172	249
	26	209	345	303	205	259	248	270	184	245
f3_4353	1	197	345	301	219	255	254	270	172	249
	19	209	345	301	207	263	256	260	172	243

Family 12	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4275	20	205	345	313	207	255	268	264	170	239
	21	211	343	303	205	263	256	260	172	243
f1_4309	2	225	345	303	205	259	248	270	184	245
	20	205	345	313	207	255	268	264	170	239
f2_4324	2	225	345	303	205	259	248	270	184	245
	21	211	343	303	205	263	256	260	172	243
f3_4406	1	197	345	301	219	255	254	270	172	249
	20	205	345	313	207	255	268	264	170	239

Family 13	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4276	3	207	345	301	211	255	264	280	172	247
	22	205	345	311	211	263	260	260	176	241
f1_4310	1	197	345	301	219	255	254	270	172	249
	3	207	345	301	211	255	264	280	172	247
f2_4325	2	225	345	303	205	259	248	270	184	245
	3	207	345	301	211	255	264	280	172	247
f3_4407	1	197	345	301	219	255	254	270	172	249
	3	207	345	301	211	255	264	280	172	247

Family 14	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4277	9	211	345	303	213	259	244	272	168	255
	23	223	345	301	207	267	264	270	172	249
f1_4311	2	225	345	303	205	259	248	270	184	245
	9	211	345	303	213	259	244	272	168	255
f2_4326	24	197	345	301	205	259	248	270	184	245
	23	223	345	301	207	267	264	270	172	249
f3_4408	2	225	345	303	205	259	248	270	184	245
	9	211	345	303	213	259	244	272	168	255

Family 15	Hap	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sir__4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4278	5	209	345	301	205	259	248	270	184	245
	27	207	345	307	211	269	260	260	178	241
f1_4312	2	225	345	303	205	259	248	270	184	245
	27	207	345	307	211	269	260	260	178	241
f2_4327	2	225	345	303	205	259	248	270	184	245
	5	209	345	301	205	259	248	270	184	245
f3_4354	2	225	345	303	205	259	248	270	184	245
	27	207	345	307	211	269	260	260	178	241

3.2 Intra-MHC Recombinant

The only foal that has a unique color, deep purple, is foal number 4326 in family 14. This foal with its haplotype is shown in **Figure 3.1** and **Table 3.3**. This haplotype (haplotype 24) formed by a recombination event within the stallion's MHC class III region. The two closest markers to the event were the ABGe_9030 and ABGe_9019, each within the region. The location of these markers can be found on the map of the MHC along with the recombination region (**Figure 3.2**). A table of the microsatellite fragment lengths for the sire, mare, and recombinant horse shows how the recombination event can be detected by this molecular technique (**Table 3.6**)

Table 3.6 Intra-MHC recombination detected by microsatellite testing. This produced a new haplotype (24) in the foal (4326, Family 14).

	Haplotype	Cor110	TAMU_305_93	ABGe_9019	ABGe_9030	TKY3324	Cor112	Cor113	UM011	Cor114
sire_4279	1	197	345	301	219	255	254	270	172	249
	2	225	345	303	205	259	248	270	184	245
dam_4277	9	211	345	303	213	259	244	272	168	255
	23	223	345	301	207	267	264	270	172	249
foal_4326	27	197	345	301	205	259	248	270	184	245
	23	223	345	301	207	267	264	270	172	249

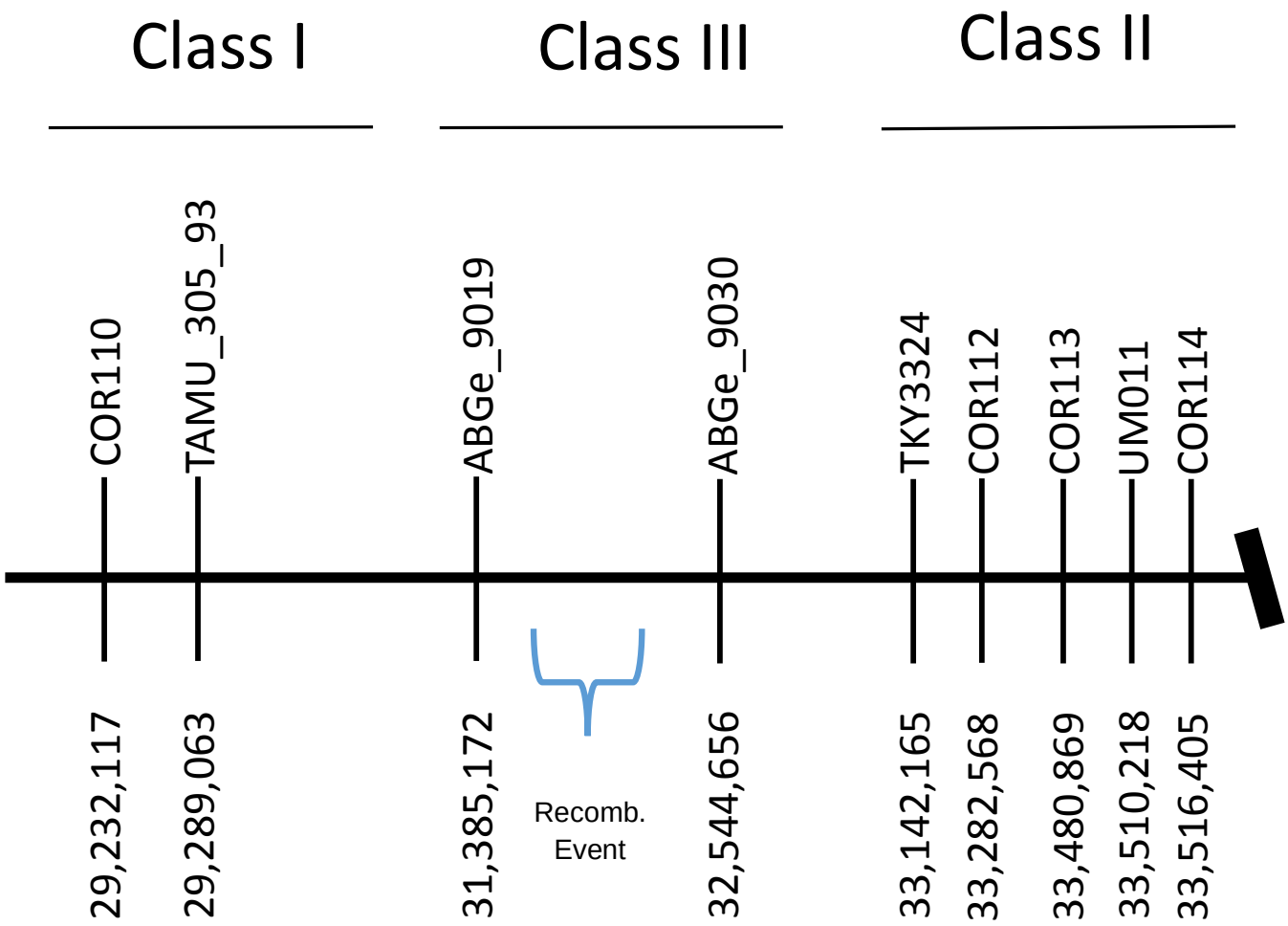

Figure 3.2 Map of Major Histocompatibility Complex microsatellite loci and location of recombination event in foal 4326. Microsatellite base pair location on ECA20 indicated.

3.3 Segregation Distortion

The offspring inherited the sire's haplotype 2 (light blue) significantly more than haplotype 1 (yellow). Fourteen foals received haplotype 1 from the sire and thirty foals received haplotype 2; the recombinant foal is excluded from this data. This segregation distortion can be seen in the lineage web by observing the majority of light blue haplotypes in the outer layer of horses (**Figure 3.1**). A chi-squared test was used to determine whether this segregation distortion was significantly different compared to the assumed inheritance of the haplotypes which would be an even 1:1 ratio based on Mendelian inheritance (**Table 3.7**). This test produced a p-value of 0.016 which indicates significant difference and concluded the haplotypes were not evenly inherited.

For comparison, I used a chi-squared test on the segregation of the stallion's sex chromosome. This tested the observed number of sexes seen in the foals to the assumed even ratio of 1 male:1 female and found that the distribution was not significantly different based on the p-value of 0.456 (**Table 3.8**).

Table 3.7 Chi-squared test for MHC haplotype segregation in the sire, assuming even segregation. The sire transmitted haplotype 2 significantly more than haplotype 1 (n=44, p=0.016, test statistic=5.82)

	Observed	Expected
Haplotype 1	14	22
Haplotype 2	30	22

Table 3.8 Chi-squared test for sex segregation in the sire, assuming even segregation. The sire did not transmit one of his sex chromosomes significantly more than the other one (n=45, p=0.456, test statistic=0.556)

	Observed	Expected
Male	25	22.5
Female	20	22.5

Chapter 4: Discussion

4.1 Haplotype Identification

The main goal of this project was to characterize haplotypes of the MHC in the herd of Icelandic horses. This goal was completed successfully with all nine of the microsatellite markers. The haplotypes were able to be phased and conclusive alleles were determined. The results showed a large degree of genetic variation in the MHC of the Icelandic herd compared to other well studied horse breeds. This can be demonstrated by the high number of unique haplotypes, 26, compared to total haplotypes in the mares and stallion, 32 (**Table 3.3**). Previous studies using similar techniques on other breeds have shown less diversity within the MHC. One paper found nineteen haplotypes in both Standardbred and Thoroughbreds in a sample size around one hundred horses for each breed (Tseng *et al.* 2010).

Looking at the diversity of alleles, the Icelandic horses showed a very similar number of alleles at each of the MHC microsatellite loci compared to studies on Thoroughbreds, Standardbreds, and mixtures of common horse breeds (Brinkmeyer-Langford *et al.* 2013). This observation indicates that even though the diversity of the alleles is similar the Icelandic horses are able to achieve a greater amount of microsatellite haplotype diversity due to different arrangements of the same alleles on the different chromosomes. In addition, other studies have found that there is a large degree of overlap in the microsatellite fragment lengths between commonly studied breeds. Comparing these fragment lengths themselves to the Icelandic herd, I have

found that the Icelandic horses have a few fragment lengths that are not found in the common breeds (Tseng *et al.* 2010). This gives evidence to the fact that the Icelandic breed diverged from the ancestors of modern breeds earlier than those breeds diverged from one another. Also, those breeds have been exposed to each other which means that some amount of gene flow must occur between them. Interbreeding is another reason for the similarities between common breeds.

The diversity in the Icelandic horses' haplotypes as well as their differing alleles is most likely a product of their history and environment. As stated in the introduction (**Chapter 1**) of this paper, the history of this breed has two key properties; the bottleneck events like the volcanic eruption and pure breeding for over a thousand years. These events would be expected to reduce genetic diversity as seen in other populations affected by bottlenecking and inbreeding (Slatkin, 1987). Other researchers argue that these conditions can also influence the population to create greater diversity through mutations and genetic mechanisms having a larger affect in a smaller population (Varvio *et al*. 1986). Our findings indicate that within the MHC of the Icelandic horse population diversity has been generated or persisted despite the relatively small founding population, large bottlenecking event and purposeful pure breeding.

A factor that may greatly influence the results of this study is the founder effect. The founder effect is important to take into account when comparing modern breeds of horses to older breeds. The modern racing breeds began from a very small number of horses. For example, all thoroughbred horses trace back to three stallions that lived about 400 years ago. This limits the genetic diversity of the breed greatly due to a very

small founding population as well as the relative youth of the breed compared to Icelandics. The Icelandic breed most likely started with a larger population of horses, compared to common racing breeds, brought over by the settlers and the breed has had thousands of years for possible diversity to come about. Although, the size of the founding population of the Icelandic breed was most likely relatively low compared to other studies with different species like *D. melanogaster* or *C. elegans.*

There are a few factors in play that could have influenced the MHC diversity that exists in the Icelandic herd, the first of which would be a recombination-like mechanism that has generated diversity within the MHC since the first founding group of horses reached Iceland. These genetic events would be able to produce the observed wide variety of haplotypes from a smaller pool of alleles due to accumulation over the long evolutionary time span. A very similar case of MHC diversity found in a small founding population is described in Doxiadis *et al.* 2013. This paper studied populations of Indian Rhesus macaques that separated from the larger Chinese population about 20 thousand years ago. This population also experienced a severe bottleneck that decreased their population by about four-fold due to geographic events such as glacial conditions and desiccation. The study found a large degree of diversity in these monkeys and the researchers equate it with recombination-like events that caused unique haplotypes to arise over time.

Another possible factor for the diversity that I found in the Icelandic MHC haplotypes would be that these unique haplotypes produced in the Icelandics are negatively selected against in more common horse populations. This would mean that some of the haplotypes that persist in the Icelandic population would not be favorable in

larger horse populations that have more diverse and more pathogenic diseases. Therefore, these less fit haplotypes would be seen in the Icelandic population but not in more commonly studied breeds. This conclusion relies on the assumption that the Icelandic environment has fewer equine pathogens. This assumption can be reinforced due to Icelandic horses' poor immune response to pathogens when brought to the United States and other countries. For example, the population of horses in Iceland are naïve to various agents, like equine influenza, strangles, and herpes abortion virus, and will likely show severe reactions to these agents (Svansson, 2004).

4.2 Intra-MHC Recombinant

The recombination event that I was able to detect within the class III region of the MHC in the foal, 4326, gives a clear example of how recombination has the ability to create unique haplotypes over time (**Figure 3.2**). It is hard to estimate the likelihood of this event due to lack of background studies of this region in Icelandic horses. Looking at other studies, the recombination rate for the human MHC has been found to be about 5.33% (Fulya *et al.* 2012) and about 0.4% in the Indian Rhesus Macaque (Doxiadis *et al.* 2013). The actual rate of MHC recombination in Icelandic horses may be close to the 2.2% recombination rate (1 out of 45 meiosis) produced by the stallion.

This specific region where the recombination event could be detected in the foal may be a hotspot for recombination. This part of the MHC between the microsatellite markers ABGe_9030 and ABGe_9019 (about 1,000,000 base pairs) has shown recombination in other tests conducted in the Antczak laboratory. Two Arabian horses

and one Standardbred horse that have been tested using a SNP array have shown crossing over between known haplotypes in this same region.

Looking closely at the microsatellite data, evidence of previous recombination events can be seen in the haplotypes. **Table 3.3** shows all of the haplotypes identified. Many of these haplotypes have the same fragment lengths for the certain class regions but have differing lengths in other regions. These data suggest that many of the haplotypes are genetically identical for those regions that have the same microsatellite fragment lengths yet differ in the other regions. This provides evidence that past recombination events have occurred to produce new haplotypes.

It is likely that the diversity in certain populations that have experienced bottlenecking, despite a small number of alleles in the founders, came about by the recombination of relatively small segments of the MHC over a long period of time. There is evidence in other species that this mechanism may be a source of genetic diversity in the MHC for populations under these circumstances as I have stated previously regarding Doxiadis *et al.* 2013.

4.3 Segregation Distortion

The possible segregation distortion in chromosome 20 of the stallion was very apparent on first glance at the lineage web (**Figure 3.1**) and the table of families showing only haplotype numbers and colors (**Table 3.3**). In the lineage web and the table the light blue haplotype (haplotype 2) is much more frequent than the yellow haplotype (haplotype 1). The chi-squared test confirmed the notion that there was a

significant difference between the observed rate that the haplotypes were inherited and the expected rate of inheritance (**Table 3.7**). The expected rate of inheritance was even for each of the two haplotypes due to traditional Mendelian rules. We compared the distribution of the MHC microsatellite haplotype segregation to the segregation in the stallion's sex chromosome in order to control for segregation distortion throughout the genome. The segregation of the sex chromosome, determined by the sex of the foals, showed no significant difference based on the chi-squared test. This result can be seen visually in **Figure 3.1** by comparing the distribution of pink and blue foals to the distribution of foals that received the yellow haplotype to the light blue haplotype. This result reinforced the possibility of true MHC segregation distortion in the stallion.

Segregation distortion in horses has been identified in a study conducted on a Standardbred stallion (Bailey, 1986). Like the Icelandic stallion, this study found segregation distortion in the MHC or chromosome 20. Compared to my study, a greater number of foals were tested. Also, the male offspring with the more frequent haplotype in Bailey, 1986 were seen to show segregation distortion in their offspring as well. This shows that further testing of the Icelandic stallion and his offspring could show results that parallel this study.

4.4 Future Studies

In regards to haplotype identification as well as recombination, going on a trip to Iceland in order to obtain more blood samples to test would significantly increase the sample size of the study. This would allow for a better estimate of diversity in the MHC of the Icelandic breed as well as the rate of recombination seen in the MHC in this

breed. The best way to do this would be to get blood from as many trios (stallion, mare, and foal) as possible. This way haplotypes could be easily phased and recombination could be detected. On the immunological side of the haplotype identification, the haplotypes that I have determined could be compared to immunological testing done on the horses to see if there is any correlation.

Another future direction would be to look further into the recombination hotspot in the class III region. Different molecular techniques could be used on the horses known to have exhibited this recombination to obtain higher resolution of the region. This could possibly show similarities in the regions where recombination occurred and give an explanation for the higher rate in this particular area.

The segregation distortion found in the stallion could be looked into further by breeding the stallion and testing his offspring. Also, his male offspring that received haplotype 2 can be bred and segregation distortion could be tested in their offspring.

Chapter 5: Literature Cited

Abbas, Abul K., Lichtman, Andrew H., Pober, Jordan S. Cellular and Molecular Immunology, 4th edition Philidelphia, PA: W.B. Saunders, 2000.

Adalsteinsson, S. (1978) Inheritance of yellow dun and blue dun in the Icelandic toelter horse. *The Journal of Heredity,* 69:146-148.

Bailey E. (1986) Segregation distortion within the equine MHC; analogy to a mouse T/t-complex trait. *Immunogenetics* 24: 225 – 9.

Brinkmeyer-Langford, C. L., Cai, J. J., Gill, C. A. and Skow, L. C. (2013) Microsatellite variation in the equine MHC. *Animal Genetics*, 44: 267–275.

Campana, M. G., Stock, F., Barrett, E., Benecke, N., Barker, G. W. W., Seetah, K. and Bower, M. A. (2012) Genetic stability in the Icelandic horse breed. *Animal Genetics*, 43: 447–449.

Díaz S., Giovambattista G., Peral-García P. (2005) Polymorphisms of the upstream regulatory region of the major histocompatibility complex DRB genes in domestic horses. *International Journal of Immunogenetics,* 32:91-98.

Doxiadis, G.G.M., de Groot, N., Otting, N., de Vos-Rouweler, AJM., Bolijn, MJ., Heijmans, C.M.C., de Groot, N.G., van der Wiel, M.K.H., Remarque, E.J., Vangenot, C., Nunes, J.M., Sanchez-Mazas, A., Bontrop, R.E. (2013) Haplotype diversity generated by ancient recombination-like events in the MHC of Indian rhesus macaques. *Immunogenetics,* 65:569–584.

Fraser, D.G., Bailey, E. (1996) Demonstration of three DRB loci in a domestic horse family. *Immunogenetics,* 44:441–445.

Fulya, T., Ender, A. (2012) Meiotic recombination within major histocompatibility complex of human embryos. *Immunogenetics,* 64:839-44.

Hendricks, Bonnie. (1995) *International Encyclopedia of Horse Breeds*. University of Oklahoma Press. p. 232. ISBN 978-0-8061-3884-8.

Hopman, T. J., Han, E. B., Story, M. R., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., Antczak, D. F. (1999) Equine dinucleotide repeat loci COR001-COR020. *Animal Genetics*, 30: 225–226.

Klumplerova, M., Vychodilova, L., Bobrova, O., Cvanova, M., Futas, J., Janova, E., Vyskocil, M., Vrtkova, I., Putnova, L, Dusek, L., Marti, E., Horin, P. (2013) Major histocompatibility complex and other allergy-related candidate genes associated with insect bite hypersensitivity in Icelandic horses. *Molecular Biology Reports,* 40:3333-40.

Lam, T., Shen, M., Chia, J., Chan, S.H., Ren, E.C. (2013), Population-specific recombination sites within the human MHC region. *Heredity,* 111: 131-138.

Ruth, L. S., Hopman, T. J., Schug, Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R. and Antczak, D. F. (1999) Equine dinucleotide repeat loci COR041-COR060. *Animal Genetics*, 30: 320–32.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science,* 236:787-792.

Svansson, V. (2004) Viral infections in horses in Iceland. In: Proceedings of the International Symposium on Diseases of the Icelandic Horse*, International Veterinary Information Service*, 30–31.

Torfason, E.G., Thorsteinsdottir, L., Torsteinsdottir, S., Svansson, V. (2008) Study of equid herpesviruses 2 and 5 in Iceland with a type-specific polymerase chain reaction. *Research in Veterinary Science*, 85:605–611.

Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., Miura, N., Choi-Miura, N., Tomito, M. (2001) Population study and validation of paternity testing for Thoroughbred horses by 15 microsatellite loci. *The Journal of Veterinary Medical Science,* 63:1191–7.

Traherne J. A. (2008) Human MHC architecture and evolution: implications for disease association studies. *International Journal of Immunogenetics,* 35:179-92.

Tseng, C. T., Miller, D., Cassano, J., Bailey, E. and Antczak, D. F. (2010), Identification of equine major histocompatibility complex haplotypes using polymorphic microsatellites. *Animal Genetics*, 41:150–153.

Varvio SL, Chakraborty R, Nei M. (1986) Genetic variation in subdivided populations and conservation genetics. *Heredity,* 57:189-198.